EXHIBIT D

Cell, Vol. 51, 503–512, November 6, 1987, Copyright © 1987 by Cell Press

Site-Directed Mutagenesis by Gene Targeting in Mouse Embryo-Derived Stem Cells

Kirk R. Thomas and Mario R. Capecchi
Department of Biology
University of Utah
Salt Lake City, Utah 84112

Summary

We mutated, by gene targeting, the endogenous hypoxanthine phosphoribosyl transferase (HPRT) gene in mouse embryo-derived stem (ES) cells. A specialized construct of the neomycin resistance (*neo*^{r}) gene was introduced into an exon of a cloned fragment of the *Hprt* gene and used to transfect ES cells. Among the G418^r colonies, 1/1000 were also resistant to the base analog 6-thioguanine (6-TG). The G418^r, 6-TG^r cells were all shown to be *Hprt*$^-$ as the result of homologous recombination with the exogenous, *neo*^{r}-containing, *Hprt* sequences. We have compared the gene-targeting efficiencies of two classes of *neo*^{r}–*Hprt* recombinant vectors: those that replace the endogenous sequence with the exogenous sequence and those that insert the exogenous sequence into the endogenous sequence. The targeting efficiencies of both classes of vectors are strongly dependent upon the extent of homology between exogenous and endogenous sequences. The protocol described herein should be useful for targeting mutations into any gene.

Introduction

Gene targeting—the homologous recombination of DNA sequences residing in the chromosome with newly introduced DNA sequences—provides a means for systematically altering the mammalian genome (Smithies et al., 1985; Thomas et al., 1986; Thomas and Capecchi, 1986). A desired alteration would first be introduced into a cloned DNA sequence, and gene targeting would then transfer the alteration into the genome. Gene targeting should be equally effective for correcting or mutating the desired chromosomal locus.

We initiated our analysis of gene targeting in cultured mammalian cells by studying recombination between a defective gene residing in the chromosome and newly introduced plasmid DNA carrying a different mutation in the same gene. For those experiments, we first established cell lines containing a mutant neomycin resistance gene (*neo*^{r}) integrated into the genome of mouse L cells. We were then able to restore the gene via homologous recombination by injecting DNA carrying a different mutation in the *neo*^{r} gene. In the course of these experiments we uncovered two mechanisms for altering chromosomal sequences. The first involved the transfer of information, by homologous recombination, from the newly introduced DNA into the cognate chromosomal sequence (Thomas et al., 1986). The second involved inducing mutations in the homologous chromosomal sequence by what appears to entail incorrect repair of a heteroduplex formed between the newly introduced DNA and the cognate chromosomal sequence (Thomas and Capecchi, 1986). Each of the two methods has its own advantages. The transfer of information by homologous recombination allows one to mutate or correct the desired chromosomal locus in a defined manner. On the other hand, the frequency of altering chromosomal sequences by heteroduplex-induced mutagenesis promises to be higher than via homologous recombination. This could prove to be a useful method for generating a large number of random mutations in specific genes.

In this current study we have extended our analysis of gene targeting by using an endogenous gene as the target and by using embryo-derived stem (ES) cells as the recipient cell line.

The target gene is hypoxanthine phosphoribosyl transferase (*Hprt*). This gene was selected primarily for two reasons. First, the *Hprt* gene lies on the X-chromosome. Since ES cells derived from male embryos are hemizygous for *Hprt*, only a single copy of the *Hprt* gene needs to be inactivated in order to yield a selectable phenotype. Second, selection procedures have been developed for isolating *Hprt*$^-$ mutants. By far the most common pathway for cells in culture to become resistant to the base analog 6-thioguanine (6-TG) is to acquire a mutation in the *Hprt* gene (Sharp et al., 1973; Wahl et al., 1975).

ES cells were chosen for these experiments because, following inactivation of a chosen gene by gene targeting, they should provide the means to generate mice with the desired mutation. ES cells have been shown to be pluripotent in vitro and in vivo (Evans and Kaufman, 1981; Martin, 1981). When reintroduced into mouse blastocystes, these cells contribute efficiently to the formation of chimeras, including contributions to a functional germ line (Bradley et al., 1984). In addition, it has been shown recently that these cells can be manipulated in vitro without losing their capacity to generate germ-line chimeras. Following transfection with the *neo*^{r} gene and selection for G418^r, these ES cells were used to produce germ-line chimeras that stably transmitted G418^r to subsequent generations (Gossler et al., 1986; Robertson et al., 1986). HPRT-deficient mice were produced from ES cells that were either selected for spontaneous *Hprt*$^-$ mutations (Hooper et al., 1987) or selected for *Hprt*$^-$ following the random insertion of retroviral DNA into the mouse genome (Kuehn et al., 1987).

Here we describe the site-directed inactivation of the endogenous *Hprt* gene in male ES cells by gene targeting. We examine some parameters that affect the gene-targeting frequency as well as the mechanism of gene inactivation mediated by different recombinant vectors. Under optimal conditions, we find that 1/1000 cells transformed by exogenous DNA can undergo a gene-targeting event. The advantage of inactivating specific genes via gene targeting compared with random mutagenic methods such

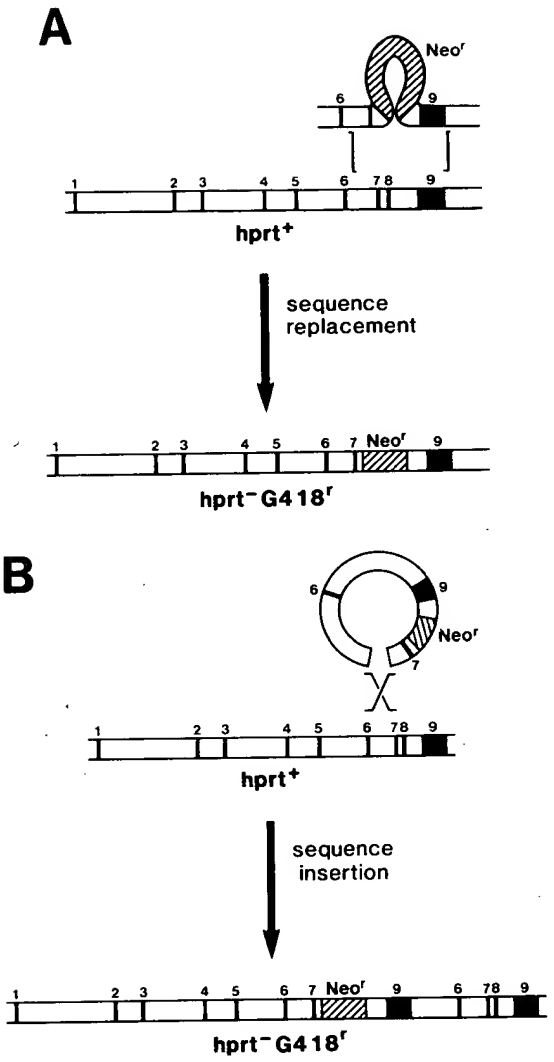

Figure 1. Disruption of the *Hprt* gene by Gene Targeting

Two schemes for gene disruption, one by sequence replacement vectors and one by sequence insertion vectors, are depicted. Vectors of both classes contain *Hprt* sequences interrupted in the eighth exon with the *neo^r* gene.

(A) Sequence replacement. Sequence replacement vectors are designed such that upon linearization, the vector *Hprt* sequences remain colinear with the endogenous sequences. Following homologous pairing between vector and genomic sequences, a recombination event replaces the genomic sequences with the vector sequences containing the *neo^r* gene.

(B) Sequence insertion. Sequence insertion vectors are designed such that the ends of the linearized vector lie adjacent to one another on the *Hprt* map. Pairing of these vectors with their genomic homolog, followed by recombination at the double strand break, results in the entire vector being inserted into the endogenous gene. This produces a duplication of a portion of the *Hprt* gene. Open boxes indicate introns; closed boxes indicate exons, numbered according to the map of Melton et al. (1984); the crosshatched box indicates the *neo^r* gene.

as chemical mutagenesis or retroviral DNA insertion is twofold. First, the nature of the mutant allele is at the discretion of the experimenter, and second, unlike random mutagenic events, the frequency of the targeting events is sufficiently high to make the procedure applicable to nonselectable genes.

Results

The *Hprt* gene encompasses over 33 kb of DNA and contains 9 exons that encode 1307 nucleotides of mRNA (Melton et al., 1984). In Figure 1 we illustrate our strategies for inactivating the *Hprt* gene. The eighth exon in a cloned fragment of *Hprt* is disrupted by inserting the *neo^r* gene. Following introduction of this DNA into ES cells, homologous recombination transfers this disruption into the endogenous *Hprt* gene, rendering the cells *neo^r–Hprt$^-$* and therefore resistant to the drug G418 and the base analog 6-TG.

Using gene targeting in yeast as a paradigm (Hinnen et al., 1978; Orr-Weaver et al., 1981), we constructed two classes of vectors that we believed would disrupt the *Hprt* gene either by replacing endogenous sequences or by inserting into the endogenous sequences. We termed these recombinant *neo^r–Hprt* vectors replacement vectors (RV) and insertion vectors (IV). The mechanism of inactivating the endogenous *Hprt* gene by these two vectors is depicted in Figures 1A and 1B. It was of interest to determine whether one or the other class of vectors was more efficient at targeting. Furthermore, since the end results using these two classes of vectors were predicted to be different (note the partial duplication of the gene in Figure 1B), each could be used to generate different mutant alleles.

Reengineering the *neo^r* Gene

In the schemes outlined in Figure 1 for site-specific mutagenesis of the *Hprt* gene, the *neo^r* gene is used both to disrupt the coding sequence of the target gene and as a tag to monitor the integration of the newly introduced DNA into the recipient genome. Effective use of the *neo^r* gene as a tag requires expression of the gene in the *Hprt* locus. In general, if the *neo^r* gene is to be used in a similar fashion to inactivate other genes, it must be expressed in as many chromosomal sites as possible.

In one of our mutagenesis schemes (Figure 1A), we require the newly added *neo^r*-containing sequences to convert the endogenous gene. We suspect that the frequency of gene conversion at the target locus may be inversely proportional to the length of nonhomology in the converting sequence. This certainly appears to be the case for intrachromosomal gene conversion (Letsou and Liskay, 1987).

Keeping the above points in mind, we have redesigned the *neo^r* gene to optimize expression in ES cells while maintaining its size at a minimum. In Figure 2 we illustrate the *neo^r* gene we have modified for this purpose. It is designated pMC1Neo. The neomycin protein coding sequence (d) is from the bacterial transposon Tn5. The promoter (b) that drives the *neo^r* gene is derived from the herpes simplex virus thymidine kinase gene (HSV-*tk*). This promoter appears to be effective in embryonal carcinoma (EC) cells (Nicolas and Berg, 1983; Rubenstein et al., 1984; Stewart et al., 1985). To increase the efficiency of the *tk* promoter, we introduced a duplication of a synthetic 65 bp fragment (a) derived from the PyF441 polyoma virus enhancer. This fragment encompasses the

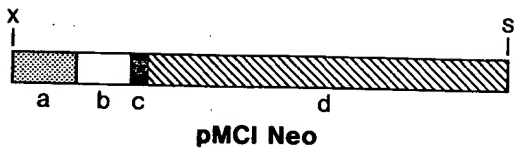

Figure 2. The *neo*^{r} Gene from pMC1Neo

The structural gene and its control elements are contained on a 1 kb cassette flanked by an XhoI site (X) and a SalI site (S) in a pUC derivative plasmid. (a) A tandem repeat of the enhancer region from the polyoma mutant PYF441 consisting of bases 5210–5274 (Fujimura et al., 1981). (b) The promoter of HSV-*tk*, from bases 92–218 (McKnight, 1980). (c) A synthetic translation initiation sequence, GCCAATATGGGATCGGCC. (d) The *neo*^{r} structural gene from Tn5, including bases 1555–2347 (Beck et al., 1982).

DNA sequence change that allows the polyoma mutant to productively infect EC cells (Linney and Donerly, 1983). Finally, because the native *neo*^{r} gene translation initiation signal is particularly unfavorable for mammalian translation, a synthetic sequence (c) was substituted using Kozak's rules as a guide (Kozak, 1986). A series of transfection experiments (data not shown) demonstrated that pMC1Neo, inserted into the eighth exon of *Hprt*, could utilize the *Hprt* poly(A) addition signal. This obviated the need to include a poly(A) addition signal in the construction.

Each of the above modifications was found to contribute additively to the transfection efficiency. The contribution of each change was assessed by introducing the different *neo*^{r} constructs into mouse fibroblasts (L cells) and mouse ES cells either by microinjection (Capecchi, 1980) or by electroporation and assaying for the yield of G418^r colonies (data not shown).

In Figure 3 we illustrate parallel experiments comparing the transfection efficiency of three *neo*^{r} vectors, pRSVNeo, pSV2Neo, and pMC1Neo, in ES cells. The DNA was introduced by electroporation. pRSVNeo contains the *neo*^{r} gene driven by the long terminal repeat from the avian Rous sarcoma virus (Hudziak et al., 1982). This promoter, with its accompanying enhancer, functions very efficiently in mouse fibroblasts (Luciw et al., 1983) but is seen here to function poorly in ES cells. pSV2Neo is an SV40 promoter-enhancer-based vector (Southern and Berg, 1982) that appears to function moderately well in ES cells. From Figure 3 it is apparent that pMC1Neo not only yields more G418^r colonies than either pRSVNeo or pSV2Neo, but also that the colonies are larger (i.e., the cells grow faster). In mouse fibroblasts, all three vectors yield G418^r colonies at comparable efficiences.

In Table 1 the transfection efficiencies of pRSVNeo, pSV2Neo, and pMC1Neo are quantitatively compared. From these data, it is apparent that pMC1Neo yields 300- to 800-fold and 25- to 50-fold more G418^r colonies than pRSVNeo or pSV2Neo, respectively. One interpretation of these results is that, since these vectors are inserting randomly into the mouse genome, the transfection efficiencies reflect the relative number of integration sites within the genome that are compatible with sufficient *neo*^{r} gene expression to yield G418^r colonies. The higher transfection efficiency of pMC1Neo may prove critical when attempting to use *neo*^{r} as a tag for targeting into genes that are either expressed at low levels, such as *Hprt*, or not at all. We have never observed G418^r colonies following mock transfections of ES cells (see Table 1).

Electroporation

We have used electroporation to introduce the *neo*^{r}–*Hprt* recombinant vectors into ES cells. The conditions for electroporation (Neumann et al., 1982; Potter et al., 1984; Chu et al., 1987) are described in Experimental Procedures. Under our optimal transfection conditions, 40%–60% of the cells survived electroporation and approximately 1/1000 surviving ES cells became G418^r. The conditions for electroporation were further chosen to yield predominantly single copy integrants. Of a dozen G418^r cell lines analyzed by Southern transfer, each contained a single copy of pMC1Neo integrated in the mouse genome (data not shown).

Recombinant Vectors

In Figure 4, we illustrate the vectors used to inactivate the *Hprt* gene. The vectors contain sequences from the 3′ portion of the mouse *Hprt* gene cloned into a pUC9 plasmid.

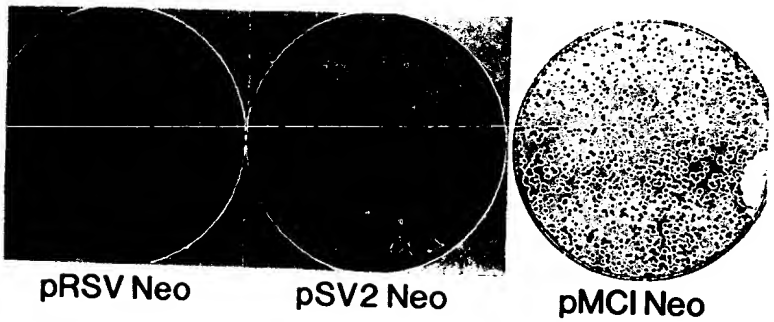

Figure 3. G418^r ES Cells Obtained by Transfection with pRSVNeo, pSV2Neo, or pMC1Neo

ES cells were transfected by electroporation with 25 μg/ml of the respective, linearized recombinant *neo*^{r} vector. The conditions for electroporation, cell culture, and selection for G418^r colonies are described in Experimental Procedures. Ten days following electroporation, the cells were fixed with cold methanol and stained with Giemsa. pMC1Neo does not contain its own poly(A) addition signal (see text). To evaluate its transfection efficiency, either a synthetic poly(A) addition signal derived from HSV-*tk*, nucleotides 1481–1530 (McKnight, 1980; Zhang et al., 1986), or a fragment from the mouse *Hprt* gene (ScaI site in exon 8 to the BglII site 3′ to the end of the gene) was added.

Table 1. Efficiency of Transfection

Vector	Exp.	No. of Cells Surviving Electroporation	No. of G418^r Colonies	Frequency of G418^r Colonies
0	1	9.0×10^7	0	0
	2	9.4×10^7	0	0
pRSVNeo	1	8.5×10^7	1.2×10^2	1.4×10^{-6}
	2	6.3×10^7	2.0×10^2	3.2×10^{-6}
pSV2Neo	1	8.7×10^7	2.1×10^3	2.4×10^{-5}
	2	9.3×10^7	4.2×10^3	4.5×10^{-5}
pMC1Neo	1	6.0×10^7	7.5×10^4	1.25×10^{-3}
	2	7.2×10^7	6.8×10^4	0.94×10^{-3}

ES cells were transfected by electroporation with 25 μg/ml of either linearized pRSVNeo, pSV2Neo, or pMC1Neo. The conditions for electroporation, cell culture, and selection for G418^r colonies are described in Experimental Procedures.

In all vectors, the 1 kb *neo^r* cassette from pMC1Neo has been inserted into the eighth exon of the *Hprt* gene. To minimize the extent of nonhomology between the endogenous *Hprt* gene and the newly introduced DNA, sequences required for growth of the recombinant vector in bacteria were removed prior to introduction into ES cells. In the process of removing these sequences, the recombinant vector is converted to linear DNA that, compared with supercoiled DNA, is a better substrate for gene targeting (Thomas et al., 1986). As discussed previously, these vectors fall into two classes, sequence replacement vectors and sequence insertion vectors, based on the predicted mode of targeting (see Figure 1).

Sequence replacement vectors were designed such that upon linearization, the vector *Hprt* sequences would remain colinear with the endogenous *Hprt* sequences. In other words, the 5′ and 3′ ends of the vector would correspond to the 5′ and 3′ extents of sequence homology with the endogenous gene (see Figure 1A). Three different sequence replacement vectors were used in this study. All of them contain a common 3′ endpoint, 2.8 kb downstream from the site of the *neo^r* gene insertion, but differ in the length of *Hprt* sequences 5′ from the insertion. The total length of *Hprt* homology in vectors pRV4.0; pRV5.4, and pRV9.1, is 4.0, 5.4, and 9.1 kb, respectively.

The alternate class of vectors, the sequence insertion vectors, were designed such that the separation of the *Hprt* vector from the pUC9 plasmids concomitantly creates a double strand break within the *Hprt* sequences. The 5′ and 3′ ends of these vectors thus lie adjacent to one another along the *Hprt* map (see Figure 1B). Two sequence insertion vectors were used in these experiments, pIV3.7 and pIV9.3. The endpoints of both linearized vectors arise from the BstEII site 1.1 kb upstream from the site of the *neo* insertion in exon 8. Both vectors contain the same 1.2 kb of *Hprt* sequences 3′ from the *neo* insertion, but differ in the length of homology at the 5′ side of the *neo* insertion.

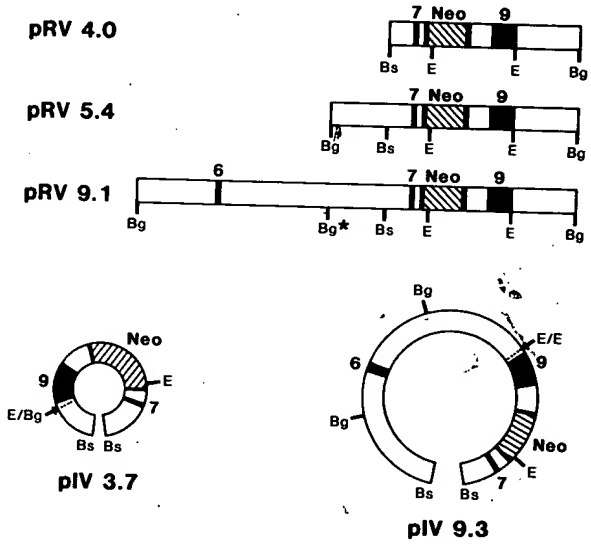

Figure 4. Targeting Vectors

The vectors were constructed as described in Experimental Procedures. The closed boxes represent *Hprt* exons, numbered according to the map of Melton et al. (1984). Open boxes represent introns and 3′-noncoding sequences. The crosshatched box represents the *neo^r* gene from pMC1Neo inserted onto the eighth exon of *Hprt*. The ends of each sequence on the diagram correspond to the site of insertion of each sequence into a pUC9-derivatized plasmid. Digestion of the plasmids with the appropriate restriction endonuclease released the *Hprt* sequences from the plasmid, creating the targeting vectors depicted above. The length of the *Hprt* sequences in each vector is as follows: pRV4.0, 4 kb; pRV5.4, 5.4 kb; pRV9.1, 9.1 kb; pIV3.7, 3.7 kb; pIV9.3, 9.3 kb. All vectors contain the poly(A) addition sequences from the *Hprt* gene. RV, replacement vector; IV, insertion vector; Bg, BglII; Bs, BstEII; E, EcoRI; the dotted line in each pIV designates the point of discontinuity of the *Hprt* gene due to the joining of the 5′ and 3′ ends during vector construction. Bg*, the internal BglII site in pRV9.1, was eliminated by cleavage with BglII and filling in with Klenow fragment and dNTPs. This permits the excision of the replacement vector using the terminal BglII sites.

Targeting with Sequence Replacement Vectors

DNAs from the three sequence replacement vectors were linearized to create the substrates shown in Figure 4 and introduced into ES cells by electroporation. Aliquots of these cells were subjected to one of three growth conditions: nonselective media, to assess the total number of cells surviving electroporation; G418 media, to assay the fraction of survivors transformed by the *neo^r*-containing vectors; and G418, 6-TG media, to select for cells simultaneously containing the *neo^r* gene but lacking a func-

Table 2. Gene Targeting Using Sequence Replacement Vectors

Vector	Exp.	No. of Cells Surviving Electroporation	No. of G418^r Colonies	No. of G418^r + 6-TG^r Colonies	G418^r + 6-TG^r / G418^r
pRV4.0	1	5.3×10^7	8.1×10^4	2	1/40,000
	2	4.3×10^7	4.3×10^4	2	1/21,500
pRV5.4	1	11.0×10^7	6.9×10^4	10	1/6,900
pRV9.1	1	7.8×10^7	3.0×10^4	32	1/950

ES cells were transfected with 25 μg/ml of linearized pRV4.0, pRV5.4, or pRV9.1. The conditions for electroporation, cell culture, selection for G418^r cell lines, and selection for G418^r, 6-TG^r cell lines are described in Experimental Procedures.

tional *Hprt* gene. As discussed below, in cells showing the G418^r, 6-TG^r phenotype, the endogenous *Hprt* gene was inactivated by the targeted replacement of the endogenous sequence with the *neo^r*-recombinant sequence.

In Table 2, we summarize the ability of the three different sequence replacement vectors to confer G418^r and G418^r, 6-TG^r upon ES cells. Although the three vectors transform ES cells to G418^r at a similar frequency (~1/1000), there is a marked difference in their ability to generate G418^r, 6-TG^r colonies. Of the G418^r cells transformed with the smallest vector, pRV4.0, only 1/40,000 to 1/20,000 showed resistance to 6-TG. However, in those cells transformed to G418^r by the largest vector, pRV9.1, 1/950 also showed the 6-TG^r phenotype. Transformation by the intermediate-sized vector, pRV5.4, gave an intermediate frequency of 6-TG resistance, with 1/7,000 G418^r colonies showing the 6-TG^r phenotype.

To show that the G418^r, 6-TG^r phenotypes were the result of gene-targeting events, the *Hprt* genes from 23 independently isolated G418^r, 6-TG^r cell lines were characterized by Southern transfer analysis. In every instance (23/23) the cells were shown to contain a single copy of the *Hprt* gene harboring the *neo^r* gene in exon 8. This result was seen in cells transformed with either pRV4.0, pRV5.4, or pRV9.1. Also, as expected, *Hprt* enzymatic activity could not be detected in these cell lines. As judged by their ability to incorporate [^{3}H]hypoxanthine into their nucleic acid, these cells contained 10^3- to 10^4-fold less activity than the parental ES cell line (data not shown).

An example of the Southern transfer analysis is shown in Figure 5A, in which the G418^r, 6-TG^r cell line EP17-2M is compared with the parental ES cell line. DNA from each line was digested with the enzymes, BglII, EcoRI, or BglII plus EcoRI, electrophoresed in agarose, and transferred to nitrocellulose paper. The paper was then hybridized with radiolabeled DNA containing 1 kb of *Hprt* sequence (Figure 5A, probe A).

As predicted from the restriction map of the cloned *Hprt* gene (see Figure 5A), digestion of the ES DNA with BglII, EcoRI, or BglII plus EcoRI isolates sequences homologous to the *Hprt* probe on fragments of lengths 5.4 kb, 9.3 kb, or 3.7 kb, respectively. The digestion pattern of the DNA from the G418^r, 6-TG^r cell line is quite different, showing fragments of 6.4 kb, 8.3 kb, and 2.7 kb. As illustrated in Figure 5C, this pattern would exist if the endogenous *Hprt* gene had been replaced by vector sequences containing the *neo^r* gene insertion. Because the *neo^r* gene insert has no BglII site, the size of the BglII fragment is increased by the size of the insert (1 kb). However, because the neo^r gene does contain an EcoRI site, its pres-

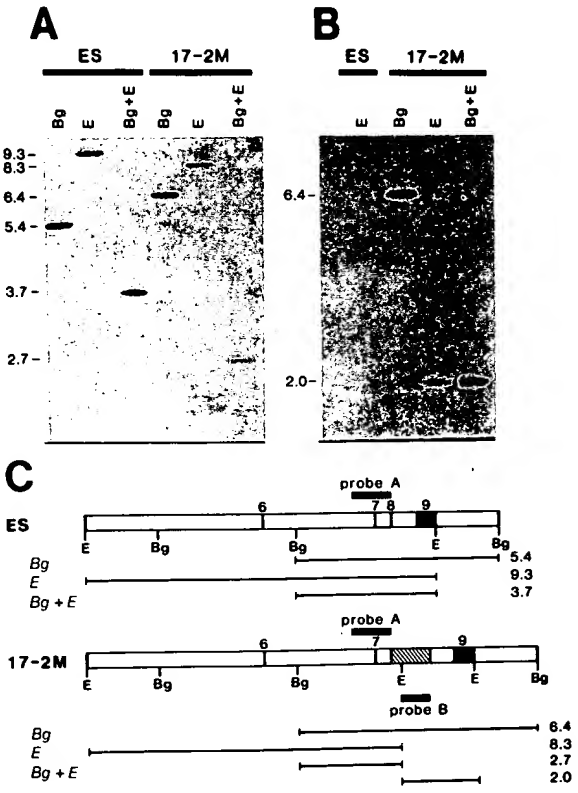

Figure 5. Southern Transfer Demonstration of Sequence Replacement

DNA was purified from each cell line and digested with restriction endonuclease. DNA (7 μg) was loaded onto an agarose gel, electrophoresed, transferred to nitrocellulose, and hybridized to ^{32}P-labeled DNA probes. ES refers to DNA from the parental, wild-type ES cell line. 17-2M refers to DNA from a G418^r, 6-TG^r cell line transformed with the replacement vector, pRV4.0. (A) Probe A was a 1 kb fragment of mouse *Hprt* DNA extending from the BstEII site in intron 6 to the ScaI site in exon 8. (B) Probe B was the 1200 bp DdeI fragment of the *neo^r* structural gene from the plasmid, pRH140 (Thomas and Capecchi, 1986). The lengths of the fragments are given in kb and were determined by the coelectrophoresis of λ and plasmid fragments of known lengths. (C) A schematic representation of the Southern transfer data. The top map represents the 3′ end of the *Hprt* gene from ES cells; the bottom represents the *Hprt* gene from the 17-2M cell line. Open boxes represent introns, closed boxes represent exons, and the crosshatched box represents the *neo^r* gene. Beneath each gene is shown the restriction fragments hybridizing to the probes. Bg, BglII; E, EcoRI.

Table 3. Gene Targeting Using Sequence Insertion Vectors

Vector	Exp.	No. of Cells Surviving Electroporation	No. of G418^r Colonies	No. of G418^r + 6-TG^r Colonies	G418^r + 6-TG^r / G418^r
pIV3.7	1	8.1×10^7	5.7×10^4	3	1/19,000
pIV9.3	1	0.74×10^7	0.42×10^4	3	1/1,400
	2	4.1×10^7	2.25×10^4	21	1/1,100

ES cells were transfected by electroporation with 25 μg/ml of linearized pIV3.7 or pIV9.3. The conditions for electroporation, cell culture, selection for G418^r cells, and selection for G418^r, 6-TG^r cells are described in Experimental Procedures.

ence introduces a new EcoRI site into the *Hprt* gene, resulting in the production of a smaller EcoRI fragment.

The interpretation is further verified when the same DNAs are hybridized to sequences from the *neo^r* gene (Figures 5B and 5C, probe B). As expected, the parental cell line contains no *neo^r* homology. The G418^r, 6-TG^r derivative does show *neo^r* homology at a site within the *Hprt* locus. Digestion of the DNA with BglII isolates the *neo^r* gene on the same 6.4 kb fragment homologous to the *Hprt* probe. Because the *neo^r* gene contains an EcoRI site at its 5′ end, digestion with this enzyme separates the *neo^r* gene from sequences homologous to the *Hprt* probe and thus creates a 2 kb fragment with *neo^r* homology.

It should be noted that the G418^r, 6-TG^r cell line EP17-2M was isolated following transformation with pRV4.0. Although this vector lacks both the EcoRI and the BglII sites 5′ to the *neo^r* insertion site (see Figure 4), the cell line EP17-2M clearly has both sites at the predicted distance from the *neo^r* gene. Such a positioning of two restriction sites is best explained by a targeted recombination event.

Gene Targeting with Sequence Insertion Vectors

The two sequence insertion vectors were linearized and introduced into ES cells by electroporation. These cells were then scored for total survivors, G418^r survivors, and G418^r, 6-TG^r survivors. The results of these experiments are summarized in Table 3. The two vectors were equally competent in the ability to confer G418^r resistance upon ES cells, but differed markedly in their ability to generate G418^r, 6-TG^r colonies. Whereas the smaller vector, pIV3.7, generates 6-TG^r cells at a frequency of 1/20,000 G418^r cells, the larger, pIV9.3, induces 6-TG^r resistance at a frequency of 1/1,100 to 1/1,400 G418^r cells. In all cases, the G418^r, 6-TG^r cells contained targeted mutations of their *Hprt* loci.

To show that gene targeting was responsible for generating the G418^r, 6-TG^r phenotype, we analyzed by Southern transfer analysis 12 cell lines transformed by pIV9.3. Unlike the case of the sequence replacement vectors in which all *Hprt* mutations were caused by the same type of event, inactivation of the *Hprt* gene by sequence insertion vectors occurred by two mechanisms. The majority of targeting events caused by pIV9.3 (9/12) were due to the insertion of the entire vector into the endogenous *Hprt* locus. The remaining targeting events were sequence replacements, resembling those events induced by the sequence replacement vectors. Examples of each event are shown in Figures 6 and 7.

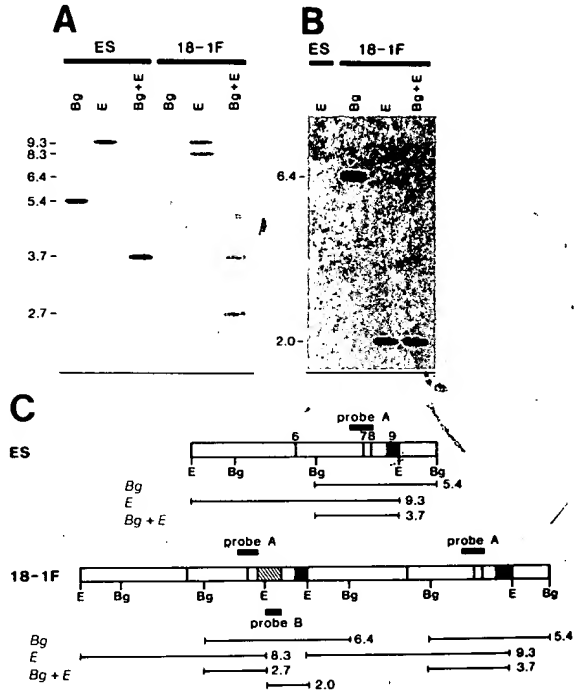

Figure 6. Southern Transfer Demonstration of Sequence Insertion

Analysis was as described in Figure 5. ES is the parental cell line. 18-1F is a G418^r, 6-TG^r cell line transformed with the insertion vector pIV9.3. (A) Hybridization with probe A containing 1 kb of *Hprt* sequences from the BstEII site in intron 6 to the ScaI site in exon 8. (B) Hybridization with probe B, the *neo^r* gene. (C) A schematic representation of the data. The top map represents *Hprt* sequences from the ES cell line. The bottom map represents sequences from the cell line 18-1F. The observed restriction fragments and their lengths are shown beneath each map. Bg, BglII; E, EcoRI.

In Figure 6 we show the Southern transfer pattern of cell line EP18-IF, a G418^r, 6-TG^r cell line transformed with pIV9.3. DNA from this cell line and DNA from the parental ES line were digested with BglII, EcoRI, or BglII plus EcoRI and probed with labeled *Hprt* sequences (Figure 6A, probe A). DNA from the parental cell line shows the 5.4, 9.3, and 3.7 kb fragments diagnostic of the wild-type *Hprt* gene. DNA from the G418^r, 6-TG^r cell line contains these same fragments, but also contains fragment of 6.4, 8.3, and 2.7 kb. These later fragments are characteristic of

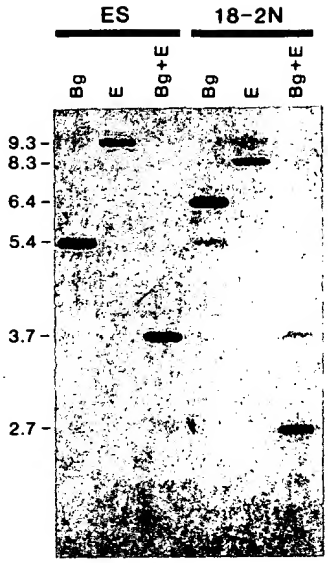

Figure 7. Southern Transfer Demonstration of Sequence Replacement Induced by a Sequence Insertion Vector

Analysis was as described in Figure 5. ES is the parental cell line; 18-2N is a G418^r, 6-TG^r cell line transformed with the insertion vector pIV9.3. The probe was the 1 kb *Hprt* fragment from the BstEII site in intron 6 to the ScaI site in exon 8 (probe A, Figure 5C). Bg, BglII; E, EcoRI. Note that the hybridization pattern is identical to that of cell line 17-2M (Figure 5A).

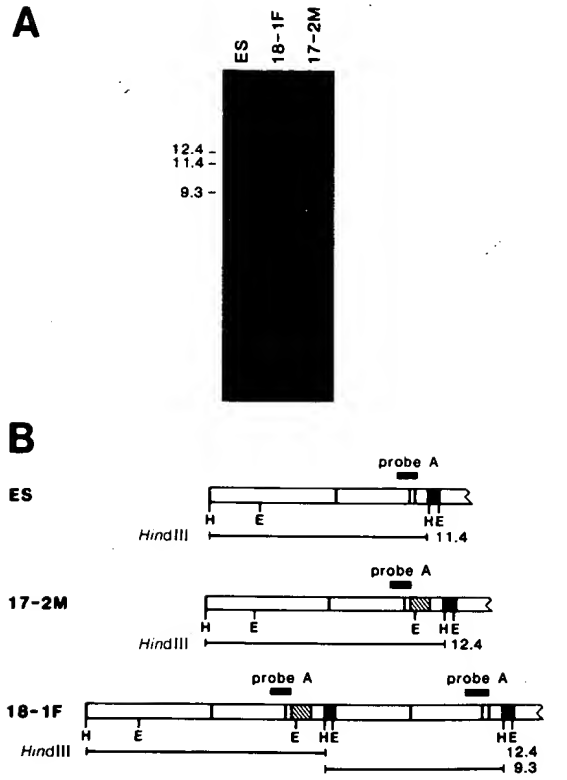

Figure 8. Southern Transfer Analysis of DNAs Digested with HindIII

DNAs were digested with HindIII and analyzed as described in Figure 5. ES is the parental cell line; 18-1F is a G418^r, 6-TG^r cell line transformed by sequence insertion; 17-2M is a G418^r, 6-TG^r cell line transformed by sequence replacement. (A) Cellular DNA was probed with *Hprt* probe A. (B) A diagram of the 3′ region of the *Hprt* genes from the three cell lines analyzed. Bg, BglII; E, EcoRI; H, HindIII.

the *Hprt* gene containing the *neo^r* gene in exon 8. One likely mechanism that would result in both fragments being recovered from the same cell is shown in Figure 6C. If the entire vector, pIV9.3 is inserted into the *Hprt* locus via homologous recombination it will cause a 9.3 kb duplication of the *Hprt* sequences. The most 5′ duplicated region will contain the *neo^r* gene, whereas the most 3′ duplicated region will contain wild-type sequences. Restriction enzyme digestions of this DNA will thus produce the hybrid configuration seen. This interpretation is further confirmed when the DNA from such a cell line is also probed with *neo^r* sequences. As shown in Figure 6B, only 1 copy of the duplicated region contains *neo^r* homology.

In 3/12 cell lines examined by Southern transfer analysis, this insertion pattern was not seen. Instead, the endogenous *Hprt* sequences appeared to have been replaced by the vector sequences containing the *neo^r* insert. An example of one such cell line, EP18-2N, transformed by pIV9.3, is shown in Figure 7. DNA from this cell line and DNA from the parental ES line were digested with BglII, EcoRI, or BglII plus EcoRI and probed with *Hprt* sequences (Figure 7). The restriction pattern generated from this DNA is indistinguishable from that generated by digestion of DNA from cell line EP17-2M (see Figure 5A), a cell line transformed by a sequence replacement vector. Thus, sequence insertion vectors are also substrates for the sequence replacement reaction.

In examining the DNA products of the sequence replacement reactions, a low level of *Hprt* sequences lacking the *neo^r* sequences is detected (see Figure 5A; Figure 7). These endogenous-length *Hprt* sequences come from the feeder cells upon which the ES cells are grown. The feeder cells are nonproliferating due to pretreatment with mitomycin C and represent a minor component of the total number of cells on the plate. However, the presence of such contaminating *Hprt* sequences presents a problem in our analysis of the sequence insertion events. In these latter events, the *neo^r*-containing vector sequences lie adjacent to the endogenous sequences, and hybridization analysis revealed the presence of these two *Hprt* sequences in equal proportions (Figure 6A). We felt compelled to demonstrate that the endogenous-length *Hprt* copy detected in this Southern transfer was in fact adjacent to the inserted copy and not the result of feeder cell contamination.

To do this, DNA was digested with the restriction endonuclease HindIII and probed with *Hprt* sequence. Such a digestion permits a distinction to be made between a single copy endogenous sequence and a sequence adjacent to the inserted vector. As shown in Figure 8, the endogenous *Hprt* sequences, represented by the parental ES cell line, are contained on an 11.4 kb fragment. Se-

quences containing a single copy of the *Hprt* gene disrupted by the *neo*^{r} gene, from cell line EP17-2M, are isolated on a 12.4 kb band. Digestion of DNA from the cell line EP18-1F, transformed by a sequence insertion vector, gives two *Hprt* fragments of equal intensity, one of 12.4 kb and one of 9.3 kb. This digestion pattern is quite consistent with a sequence insertion event.

Discussion

We have analyzed 38 independent G418^r, 6-TG^r ES cell lines. Each of these cell lines was shown to have arisen from inactivation of the endogenous *Hprt* gene by homologous recombination with the introduced *neo*^{r}*–Hprt* fragment. Spontaneous formation of G418^r, 6-TG^r cells was not detected (i.e., occurs at less than 1/10^9 cells per generation).

None of the G418^r, 6-TG^r cell lines contained extraneous copies of the *neo*^{r}*–Hprt* recombinant vector integrated randomly within the ES genome. This greatly simplified the analysis of the targeting events in these cell lines and permitted an unambiguous interpretation of the results. The absence of extraneous copies of the input vector in the targeted cells should also simplify interpretation of additional studies that will entail establishing a correlation between the inactivation of specific genes with the resultant phenotypes.

Under our optimal conditions, we have observed a gene-targeting frequency, relative to the frequency of random integration of the input vector, of 1/1000. The parameters that we believe influenced the success of these experiments include using a *neo*^{r} gene that is efficiently expressed in ES cells, maintaining the size of the *neo*^{r} gene at a minimum, using extensive homology between the homing sequence and the target sequence, and removing, prior to transfection, unnecessary and nonhomologous sequences from the input vector.

This gene-targeting frequency is sufficiently high to be used for inactivating nonselectable genes. Direct screening, by Southern transfer analysis, for a gene-targeting event among 1000 candidate cell lines would not be exorbitant. Furthermore, gene-targeting enrichment procedures could be added to the protocol for using the *neo*^{r} gene as a transfection tag. For example, a *neo*^{r} gene lacking an enhancer or a poly(A) addition signal could be positioned within the homing sequence in such a way that homologous recombination with the target gene would juxtapose the defective *neo*^{r} gene with the sequences required for effective expression. Random integration of the same vector into the recipient genome would not normally bring the required sequence sufficiently near the *neo*^{r} gene to yield G418^r colonies. Pilot experiments testing such procedures indicate that enrichment of several hundred fold for gene targeting compared with random integration should be attainable (unpublished results).

The gene-targeting frequency was observed to be very sensitive to the extent of homology between the exogenous and cognate endogenous sequence. A 2-fold increase in homology increased the gene-targeting frequency by 20-fold. Further increases in the extent of homology may increase the gene-targeting frequency even more.

We have compared two classes of *neo*^{r}*–Hprt* recombinant vectors, one that replaces endogenous sequences with exogenous sequences and another that inserts exogenous sequences into the endogenous sequence. Both classes exhibit comparable gene-targeting frequencies and are equally sensitive to the extent of homology with the endogenous target. We have termed the former sequence replacement vectors and the latter sequence insertion vectors. In 23/23 G418^r, 6-TG^r cell lines obtained by introducing the replacement vector, the endogenous *Hprt* gene was inactivated by sequence replacement. Of the 12 G418^r, 6-TG^r cell lines obtained with the insertion vector pIV9.3, 9 resulted from sequence insertion. In the remaining 3 the *Hprt* gene was inactivated by sequence replacement. The latter may result from a crossover occurring at points within the vector sequences rather than at both termini (Szostak et al., 1983). Though insertion vectors mediate mutagenesis via two pathways, they target predominantly by inserting into the endogenous gene.

The insertion vectors are technically more difficult to build. On the other hand, they may provide the means for generating a wider spectrum of mutant alleles. For example, by placing the *neo*^{r} gene in the 3′-untranslated sequence, it can still be used as a transfection tag. In such a vector, the *neo*^{r} gene could be linked to a wide spectrum of mutations, including point mutations, small insertions, or small deletions, in upstream exons. In the process of insertion of the *neo*^{r} vector, these mutations would be concomitantly transferred into the endogenous gene.

When we initiated these experiments we had two concerns about using the *Hprt* gene as our target: it is expressed at a low level in ES cells, and it contains many repetitive DNA sequences. As in most cells, HPRT protein represents approximately 1/5,000 of the soluble protein (Hughes et al., 1975). Furthermore, repetitive DNA sequences are dispersed throughout both the *neo*^{r}*–Hprt* recombinant vector and the *Hprt* gene. In fact, it is not a simple task to identify a suitable probe from the *Hprt* locus for Southern transfer analysis. The success obtained in targeting to the *Hprt* gene despite these handicaps may indicate that in the future we need not be so concerned with these parameters negatively influencing the gene-targeting frequency.

Conclusion

We have demonstrated that we can inactivate by gene targeting a specific locus in the mouse genome. The protocol we have developed to inactivate the endogenous *Hprt* gene should be adaptable to other genes as well. We have also shown that ES cells are a suitable host for gene-targeting experiments. It is hoped that this combination of using ES cells as the recipient cell line and site-specific mutagenesis achieved by gene targeting will provide the means for generating mice of any desired genotype. An advantage of this scenario is that the first generation chimera will usually be heterozygous for the targeted mutation and that subsequent breeding can be used to gener-

ate the homozygous animal. Thus, only one of the two loci need be inactivated, and recessive lethals can be maintained as heterozygotes. If successful, this technology will be used in the future to dissect the developmental pathway of the mouse as well as to generate mouse models for human genetic diseases.

Experimental Procedures

Vector Construction

Hprt sequences were isolated from a λ, Charon 4A, library containing a partial EcoRI-digest of DNA from a mouse ARK cell line (the library was provided by Doug Foster, Ohio State University). The library was screened with a human cDNA *Hprt* probe (courtesy of C. Thomas Caskey). A recombinant phage containing the 9.3 EcoRI fragment encoding *Hprt* exons 6–9, as well as the 2.2 and 1.0 kb fragments 3′ of the *Hprt* gene, was isolated. The 9.3 kb and 2.2 kb fragments were subcloned into pUC9 and converted by standard cloning methods into the targeting vectors. To introduce the *neo*^{r} gene into *Hprt* exon 8, an 8 bp XhoI linker (New England Biolabs) was ligated into the ScaI site in exon 8.

The *neo*^{r} vector pMC1Neo was created by the sequential ligation of its four functional domains (see Figure 2) into pUC9. The polyoma enhancer sequences were chemically synthesized on an Applied Biosystems model 380B DNA synthesizer and were flanked with XhoI (5′ end) and SalI (3′ end) restriction sites. To create the enhancer dimer used in pMC1Neo, the monomer units were ligated in vitro and the dimer was purified from a polyacrylamide gel. The HSV-*tk* promoter sequence from bases 92 to 121 was chemically synthesized and ligated in vitro to bases 122–218, isolated as an EcoRI–PstI fragment from the HSV-*tk* gene. The translational start sequence was synthesized chemically. The *neo*^{r} gene was derived from the transposon Tn5. The structure of pMC1Neo was confirmed by DNA sequence analysis. pMC1Neo was designed such that the *neo*^{r} gene and all its control elements could be removed as a 1 kb unit following digestion by XhoI and SalI and thus inserted into the XhoI site in the *Hprt* gene of the various targeting vectors.

Southern transfer analysis was performed as described previously (Thomas et al., 1986).

Isolation and Culturing of ES Cells

ES cells were isolated from C57B1/6 blastocysts as described by Evans and Kaufman (1981) except that primary embryonic fibroblasts (Doetschman et al., 1985) were used as feeders rather than STO cells. Briefly, 2.5 days postpregnancy mice were ovariectomized, and delayed blastocysts were recovered 4–6 days later. The blastocysts were cultured on mitomycin C–inactivated primary embryonic fibroblasts. After blastocyst attachment and the outgrowth of the trophectoderm, the ICM-derived clump was picked and dispersed by trypsin into clumps of 3–4 cells and put onto new feeders. All culturing was carried out in DMEM plus 20% FCS and 10^{-4} M β-mercaptoethanol. The cultures were examined daily. After 6–7 days in culture, colonies that still resembled ES cells were picked, dispersed into single cells, and replated on feeders. Those cell lines that retained the morphology and growth characteristic of ES cells were tested for pluripotency in vitro. These cell lines were maintained on feeders and transferred every 2–3 days. For comparative purposes we have also used ES cell lines kindly provided by Martin Evans and Gail Martin. Cell lines from all three sources yielded targeted G418^r, 6-TG^r colonies at comparable frequencies. The G418^r, 6-TG^r cell lines are morphologically indistinguishable from the parental ES cells and retained their pluripotency in vitro (i.e., differentiate when grown on petri plates in the absence of a feeder layer and form embroid bodies when grown in suspension).

Electroporation and Isolation of G418^r and G418^r, 6-TG^r Cell Lines

DNA was introduced into the ES cells by electroporation using the Promega Biotech X Cell 2000. Rapidly growing cells were trypsinized, washed in DMEM, counted, and resuspended in buffer containing 20 mM HEPES (pH 7.0), 0.37 mM NaCl, 5 mM KCl, 0.7 mM Na_2HPO_4, 6 mM dextrose, and 0.1 mM β-mercaptoethanol. Just prior to electroporation, the linearized recombinant vector was added. The cells were then exposed to a single, 625 V/cm pulse at room temperature, allowed to remain in the buffer for 10 min, and plated onto feeder cells. For every experiment, aliquots of cells were removed before and after electroporation to measure colony-forming units; 40%–60% of the cells survived electroporation.

In a typical experiment 10^7 cells per vial were transfected by electroporation with 25 μg/ml of linearized vector. Aliquots of cells were then subjected to one of three growth conditions: nonselective media, to evaluate the number of cells surviving electroporation; G418 media, to determine the fraction of survivors transformed by the *neo*^{r} vector; and G418, 6-TG media, to select cells that had simultaneously acquired a *neo*^{r} gene and lost a functional *Hprt* gene. For these experiments the ES cells were grown on mitomycin C–inactivated STO cells (obtained from Alan Bradley). To ensure inactivation, the STO cells were treated with 10 μg/ml mitomycin C for 4 hr. Survival was less than $1/10^9$ cells.

ES cells that were to be grown on nonselective medium were diluted 4×10^4-fold and 8×10^4-fold prior to plating onto 100 mm dishes containing the feeder cells. Cells that were to be subjected to growth on G418 medium or G418, 6-TG medium were diluted 200-fold and 26-fold, respectively, before plating onto feeders. To allow for expression of the *neo*^{r} gene, the cells were first plated in nonselective medium and, 48 hr later, were transferred to G418-containing medium (250 μg/ml). To allow for the decay of the endogenous HPRT activity, the cells that would eventually be subjected to G418, 6-TG selection were first plated in nonselective medium. Two days later they were transferred to G418 medium, and 5 days after electroporation they were transferred to G418, 6-TG (1 μg/ml) medium.

At each transfer, the cells were trypsinized and placed on a new feeder plate in their respective medium. It is necessary to disperse the cells prior to subjecting them to selection because ES cells grow in tight clumps and cross-feed extensively. During this period of selection, the cells are dividing and it is necessary, for quantitative analysis, to keep track of the number of cell divisions. For this purpose, aliquots of cells from the same experiment were grown in nonselective medium, subjected to the same transfer protocol, and used to measure cell proliferation during this period. An outcome of the above protocol is that if a single targeting event occurs during the time of electroporation, one of the G418, 6-TG plates will yield a burst of 32–64 G418^r, 6-TG^r colonies; the rest of the plates will contain no colonies. This is scored as a single event. Each of the G418^r, 6-TG^r cell lines that we obtained came from such individual bursts, indicating that the targeting events occurred at the time of electroporation. The above scoring procedure underestimates the gene-targeting frequency since a burst of G418^r, 6-TG^r colonies on a given plate may have arisen from more than one event. For example, if the Poisson function was used to correct the pRV9.1 data shown in Table 2 for the expected number of plates that resulted from two events, then the gene-targeting frequency would be 1/800 G418^r colonies rather than 1/950 G418^r colonies.

Acknowledgments

We would like to thank Laurie Fraser and Susan Tomowski for their excellent expert technical assistance. Without their seemingly endless labor this project would never have met fruition. We would also like to thank Alan Bradley, Elizabeth Robertson, Martin Evans, and Gail Martin for their guidance in the isolation and maintenance of ES cells. M. R. C. and Laurie Fraser want especially to acknowledge a memorable visit to Cambridge, Christmas of '85 where we gained hands-on experience with the ES system. We thank Doug Foster for providing us with the λ Charon 4A genomic library. This work was in part supported by the Cystic Fibrosis Foundation, and K. R. T. is a recipient of a Cystic Fibrosis Foundation Fellowship.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received August 12, 1987.

References

Beck, E., Ludwig, G., Auserwald, E. A., Reiss, B., and Schaller, H. (1982). Nucleotide sequence and exact localization of the neomycin

phosphotransferase gene from transposon Tn5. Gene *19*, 327–336.

Bradley, A., Evans, M., Kaufman, M. H., and Robertson, E. (1984). Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. Nature *309*, 255–256.

Capecchi, M. R. (1980). High efficiency transformation by direct microinjection of DNA into cultured mammalian cells. Cell *22*, 479–488.

Chu, G., Hayakawa, H., and Berg, P. (1987). Electroporation for the efficient transfection of mammalian cells with DNA. Nucl. Acids Res. *15*, 1311–1326.

Doetschman, T. C., Eistetter, H., Katz, M., Schmidt, W., and Kemler, R. (1985). The *in vitro* development of blastocyst-derived embryonic stem cell lines: formation of visceral yolk sac, blood islands and myocardium. J. Embryol. Exp. Morphol. *87*, 27–45.

Evans, M. J., and Kaufman, M. H. (1981). Establishment in culture of pluripotential cells from mouse embryos. Nature *292*, 154–156.

Fujimura, F. K., Deininger, P. L., Friedmann, T., and Linney, E. (1981). Mutation near the polyoma DNA replication origin permits productive infection of F9 embryonal carcinoma cells. Cell *23*, 809–814.

Gossler, A., Doetschman, T., Korn, R., Serfling, E., and Kemler, R. (1986). Transgenesis by means of blastocyst-derived embryonic stem cell lines. Proc. Natl. Acad. Sci. USA *83*, 9065–9069.

Hinnen, A., Hicks, J. B., and Fink, G. R. (1978). Transformation of yeast. Proc. Natl. Acad. Sci. USA *75*, 1929–1933.

Hooper, M., Hardy, K., Handyside, A., Hunter, S., and Monk, M. (1987). HPRT-deficient (Lesch–Nyhan) mouse embryos derived from germline colonization by cultured cells. Nature *326*, 292–295.

Hudziak, R. M., Laski, F. A., RajBhandary, U. L., Sharp, P. A., and Capecchi, M. R. (1982). Establishment of mammalian cell lines containing multiple nonsense mutations and functional suppressor tRNA genes. Cell *31*, 137–146.

Hughes, S. H., Wahl, G. M., and Capecchi, M. R. (1975). Purification and characterization of mouse hypoxanthine guanine phosphoribosyltransferase. J. Biol. Chem. *250*, 120–126.

Kozak, M. (1986). Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. Cell *44*, 283–292.

Kuehn, M. R., Bradley, A., Robertson, E. J., and Evans, M. J. (1987). A potential animal model for Lesch–Nyhan syndrome through introduction of HPRT mutations into mice. Nature *326*, 295–298.

Letsou, A., and Liskay, M. R. (1987). Effect of the molecular nature of mutation on the efficiency of intrachromosomal gene conversion in mouse cells. Genetics, in press.

Linney, E., and Donerly, S. (1983). DNA fragments from F9 PyEC mutants increase expression of heterologous genes in transfected F9 cells. Cell *35*, 693–699.

Luciw, P. A., Bishop, J. M., Varmus, H. E., and Capecchi, M. R. (1983). Location and function of retroviral and SV40 sequences that enhance biochemical transformation after microinjection of DNA. Cell *33*, 705–716.

Martin, G. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. Proc. Natl. Acad. Sci. USA *78*, 7634–7638.

McKnight, S. (1980). The nucleotide sequence and transcript map of the herpes simplex virus thymidine kinase gene. Nucl. Acids Res. *8*, 5949–5964.

Melton, D. W., Konecki, D. S., Brennand, J., and Caskey, C. T. (1984). Structure, expression and mutation of the hypoxanthine phosphoribosyl transferase gene. Proc. Natl. Acad. Sci. USA *81*, 2147–2151.

Neumann, E., Schaefer-Ridder, M., Wang, Y., and Hofschneider, P. H. (1982). Gene transfer into mouse lyoma cells by electroporation in high electric fields. EMBO J. *1*, 841–845.

Nicolas, J. F., and Berg, P. (1983). Regulation of expression of genes transduced into embryonal carcinoma cells. In Teratocarcinoma Stem Cells, L. M. Silver, G. R. Martin, and S. Strikland, eds. (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory), pp. 469–497.

Orr-Weaver, T. L., Szostak, J. W., and Rothstein, R. J. (1981). Yeast transformation: a model system for the study of recombination. Proc. Natl. Acad. Sci. USA *78*, 6354–6358.

Potter, H., Weir, L., and Leder, P. (1984). Enhancer-dependent expression of human κ immunoglobulin genes introduced into mouse pre B lymphocytes by electroporation. Proc. Natl. Acad. Sci. USA *81*, 7161–7165.

Robertson, E., Bradley, A., Kuehn, M., and Evans, M. (1986). Germ-line transmission of genes introduced into cultured pluripotential cells by retroviral vectors. Nature *323*, 445–447.

Rubenstein, J. L. R., Nicolas, J. F., and Jacob, F. (1984). Construction of a retrovirus capable of transducing and expressing genes in multipotential embryonic cells. Proc. Natl. Acad. Sci. USA *81*, 7137–7140.

Sharp, J. D., Capecchi, N. E., and Capecchi, M. R. (1973). Altered enzymes in drug-resistant variants of mammalian tissue culture cells. Proc. Natl. Acad. Sci. USA *70*, 3145–3149.

Smithies, O., Gregg, R. G., Boggs, S. S., Koralewski, M. A., and Kucherlapati, R. S. (1985). Insertion of DNA sequences into the human chromosome β-globin locus by homologous recombination. Nature *317*, 230–234.

Southern, P. J., and Berg, P. (1982). Transformation of mammalian cells to antibiotic resistance with a bacterial gene under control of the SV40 early region promoter. J. Mol. Appl. Genet. *1*, 327–341.

Stewart, C. L., Vanek, M., and Wagner, E. F. (1985). Expression of foreign genes from retroviral vectors in mouse teratocarcinoma chimaeras. EMBO J. *4*, 3701–3709.

Szostak, J. W., Orr-Weaver, T. L., Rothstein, R. J., and Stahl, F. W. (1983). The double-strand-break repair model for recombination. Cell *33*, 25–35.

Thomas, K. R., and Capecchi, M. R. (1986). Introduction of homologous DNA sequences into mammalian cells induces mutations in the cognate gene. Nature *324*, 34–38.

Thomas, K. R., Folger, K. R., and Capecchi, M. R. (1986). High frequency targeting of genes to specific sites in the mammalian genome. Cell *44*, 419–428.

Wahl, G. M., Hughes, S. H., and Capecchi, M. R. (1975). Immunological characterization of hypoxanthine-guanine phosphoribosyl transferase mutants of mouse L cells: evidence for mutations at different loci in the hprt gene. J. Cell Physiol. *85*, 307–320.

Zhang, F., Denome, R. M., and Cole, C. N. (1986). Fine-structure analysis of the processing and polyadenylation region of the herpes simplex virus type 1 thymidine kinase gene by using linker scanning, internal deletion and insertion mutations. Mol. Cell. Biol. *6*, 4611–4623.

This Page is inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☑ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLORED OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REPERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.
As rescanning documents *will not* correct images problems checked, please do not report the problems to the IFW Image Problem Mailbox